Abstract

A system is controlled by an actor-critic based fuzzy reinforcement learning algorithm that provides instructions to a processor of the system for applying actor-critic based fuzzy reinforcement learning. The system includes a database of fuzzy-logic rules for mapping input data to output commands for modifying a system state, and a reinforcement learning algorithm for updating the fuzzy-logic rules database based on effects on the system state of the output commands mapped from the input data. The reinforcement learning algorithm is configured to converge at least one parameter of the system state to at least approximately an optimum value following multiple mapping and updating iterations. The reinforcement learning algorithm may be based on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered. The reinforcement learning algorithm may be configured to update the at least one parameter based on said update equation. The system may include a wireless transmitter.